

App store analysis

Gabriel Scozzarro

04/01/2022

1.0 Introduction

The ever-changing mobile landscape is a challenging space to navigate. . The percentage of mobile over desktop is only increasing. Android holds about 53.2% of the smartphone market, while iOS is 43%. To get more people to download your app, you need to make sure they can easily find your app. Mobile app analytics is a great way to understand the existing strategy to drive growth and retention of future user.

With million of apps around nowadays, the following data set has become very key to getting top trending apps in iOS app store. This data set contains more than 7000 Apple iOS mobile application details. The data was extracted from the iTunes Search API at the Apple Inc website.

1.1 Scope of work & Business task

This study aim to offer general insights obtained from an exploratory data analysis of the App Store data.

2.0 Data

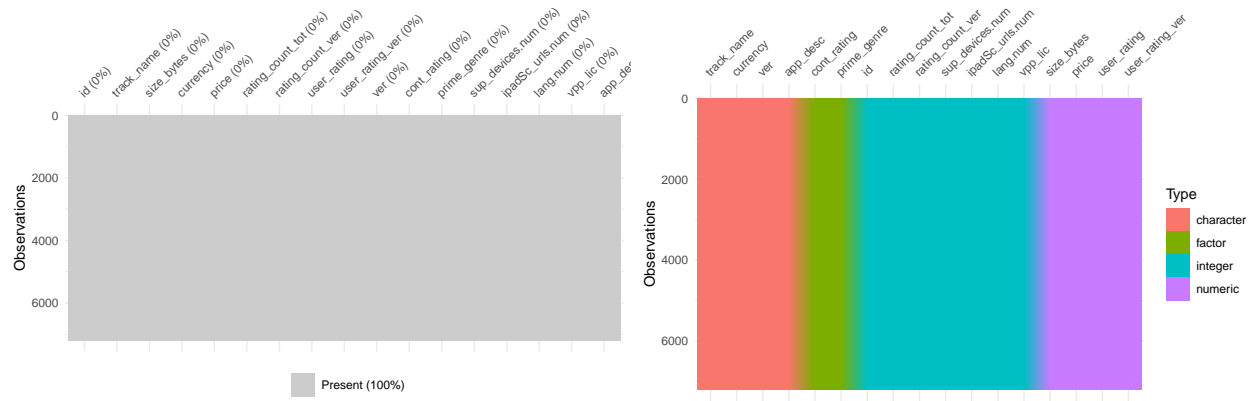
The data was extracted from the iTunes Search API at the Apple Inc website and store it on **Kaggle**. This data set is provided in csv format and has a Kaggle was already cleaned and features were drescribed for easy understanding. The data set was updated for the last time on June 2018 with data collected on July 2017 and it contains general information, description and user rate of each App.

3.0 Tools and process

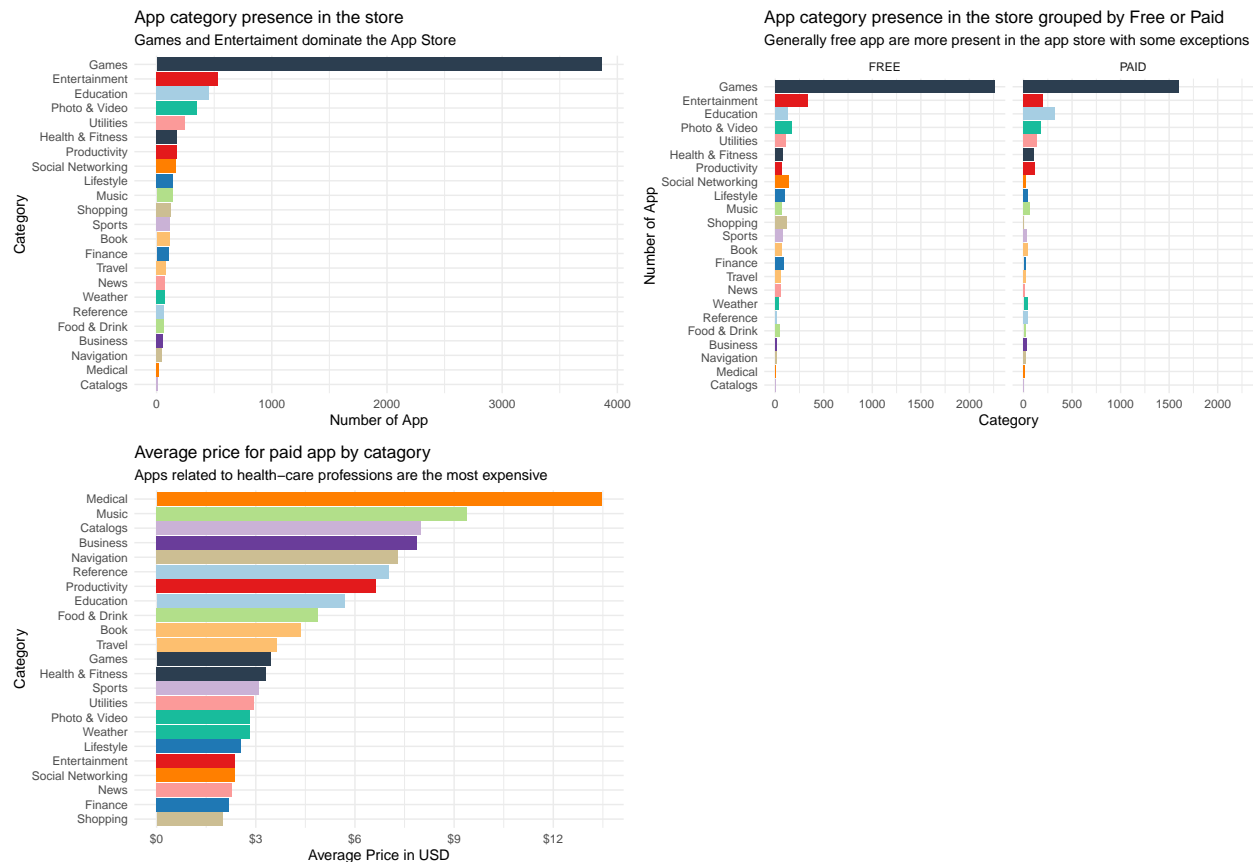
The analysis was performed using R coding language. A complete list of R packages and the source code are all avaiable on the project github repository.

4.0 Analysis

The Dataset contain information about 7197 Apps with 18 variables. 23 different categories of Apps were identified.

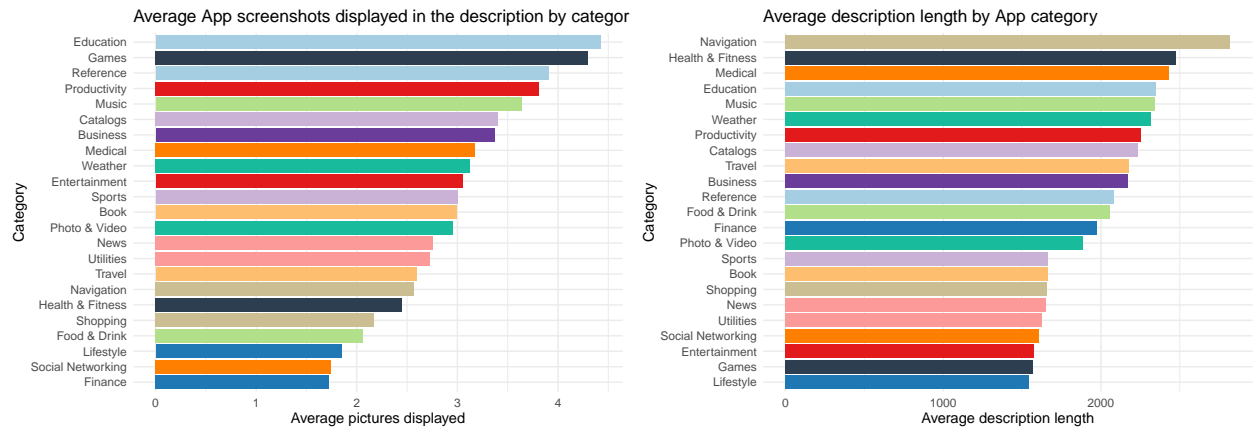


A preliminary sanity check was performed, it shows no NAs and small data type modifications were made.

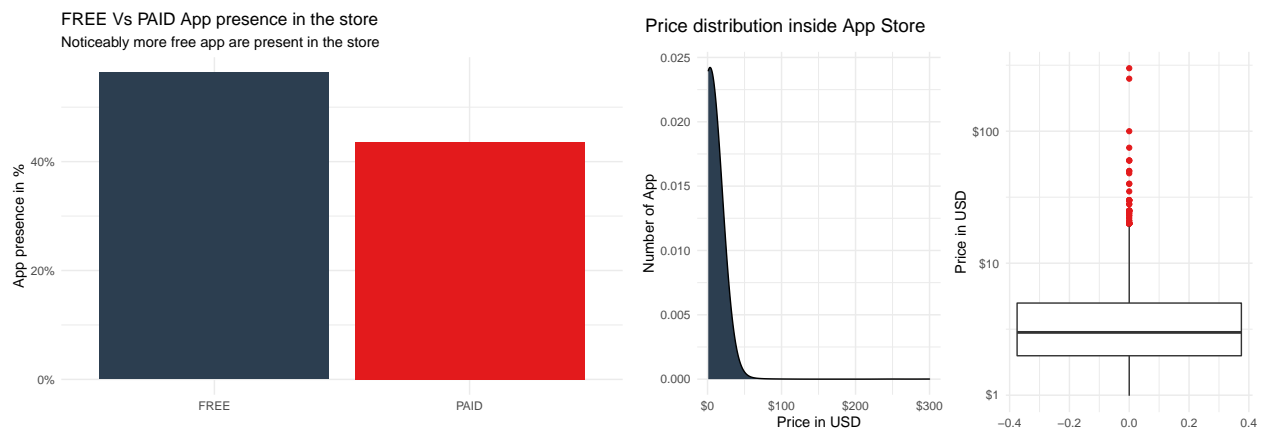


Games are the most frequent App category in the store followed by Entertainment and Education. Free Apps are noticeably more present but almost same category proportions are maintained also when we split the App store in paid App and free App.

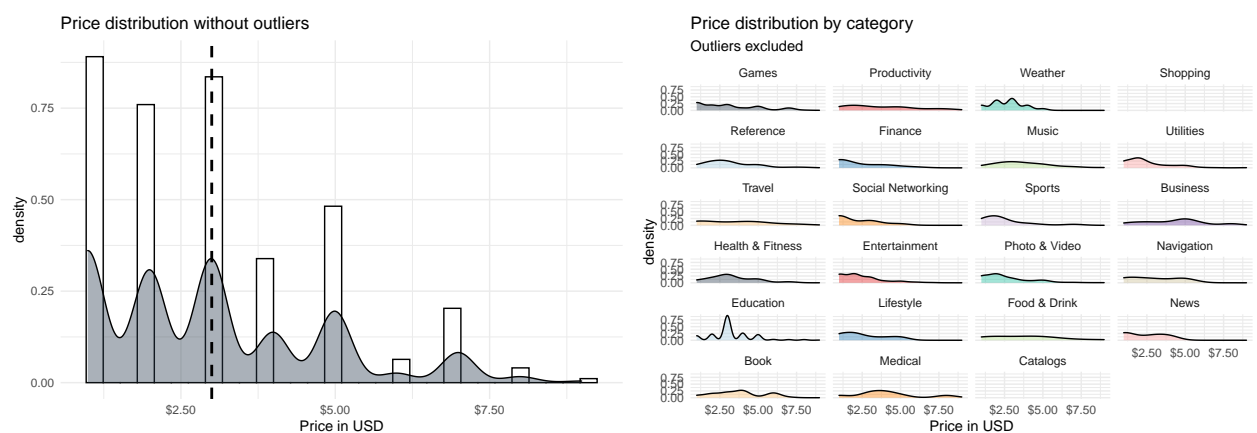
Excluding free Apps, Medical related Apps are, on average, more expensive with an average price of 13.5. Work and study related Apps has also a noticeably bigger price compared to entertaining and free time Apps.



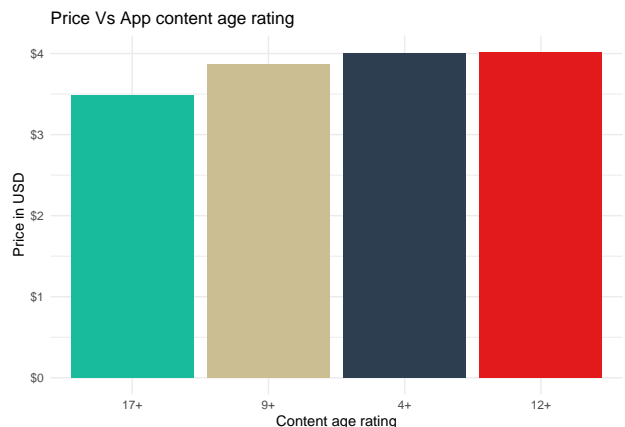
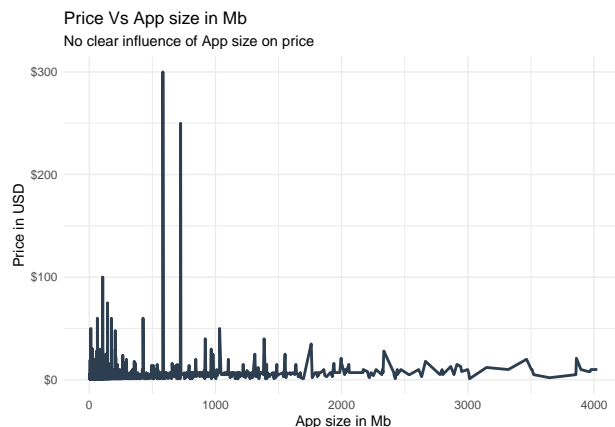
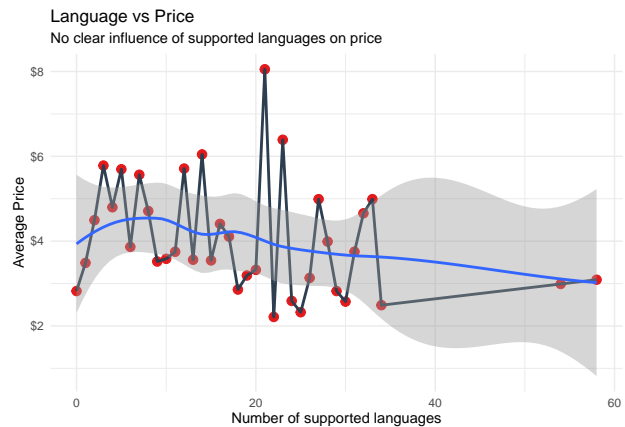
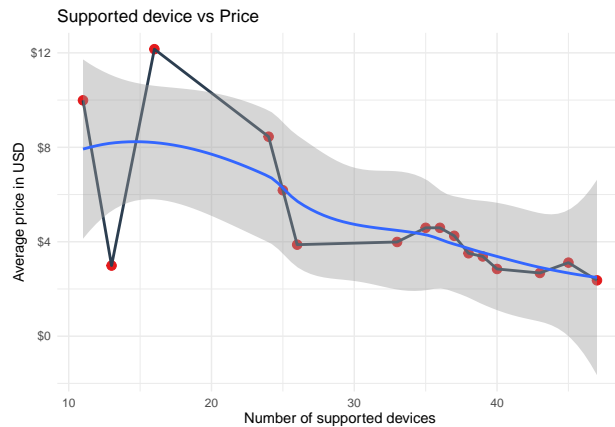
An analysis on how the different categories communicate to the customers was made. Some categories are more prone to communicate with images rather than text. For example Games on average is the second most App category for the usage of images but if we look at the other plot we can see that Games is the second last for text length.



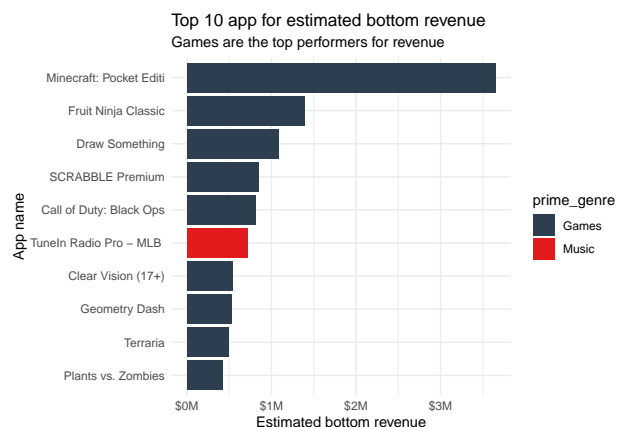
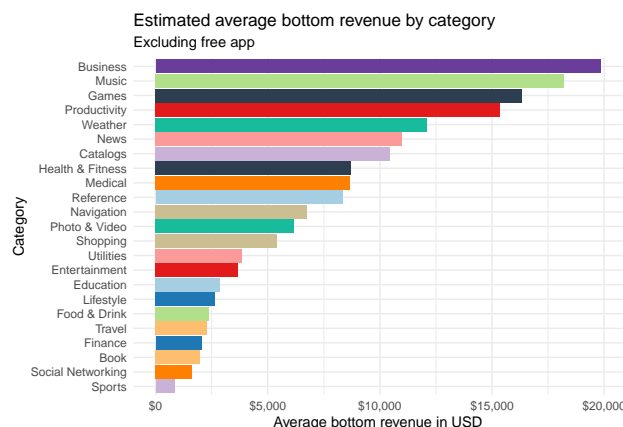
As previously mentioned there are more free App than paid ones. Free app were filtered out and a price distribution analysis was performed showing that prices, regardless the category, has a bottom margin of 0.99 and an upper margin of 299.99 with an average of 3.9552977 and standard deviation of 8.3158818. Is noticeable that prices are very skewed on the left and some really big outliers are present.



If outliers are removed the plot shows clearly how the prices inside the App store are divide in tiers. If a division in category is applied keeping filtered out free Apps and outliers some trends in pricing the Apps. (Catalogs and Shopping once filtered for outliers and free Apps has too few data points)



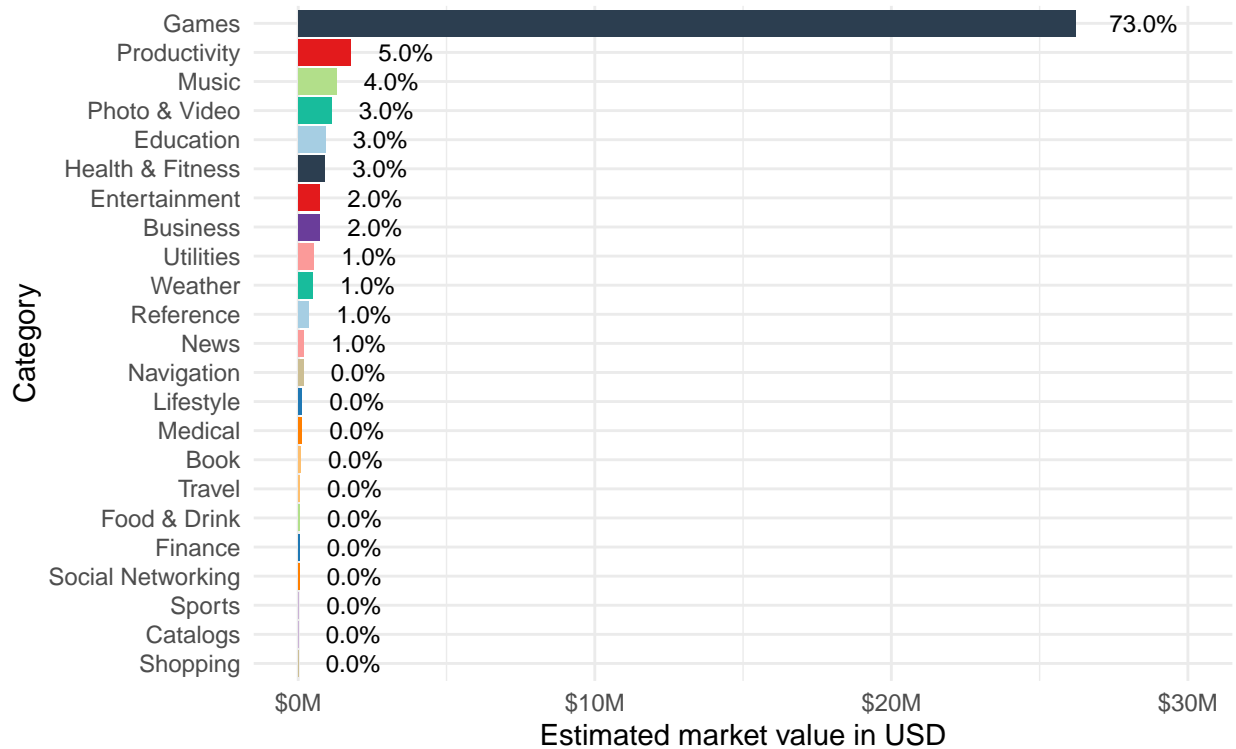
Some deeper investigation to possible relation that can influence the price of an App was performed. As the number of supported devices grows there is a trend in lowering the price which is counter intuitive if we suppose that more supported device equal more development cost. Instead, the number of supported languages and the size of the App in Mb didn't reveal an influence on the App price. The Price Vs content age rating reveal that App for adult are slightly cheaper on average.



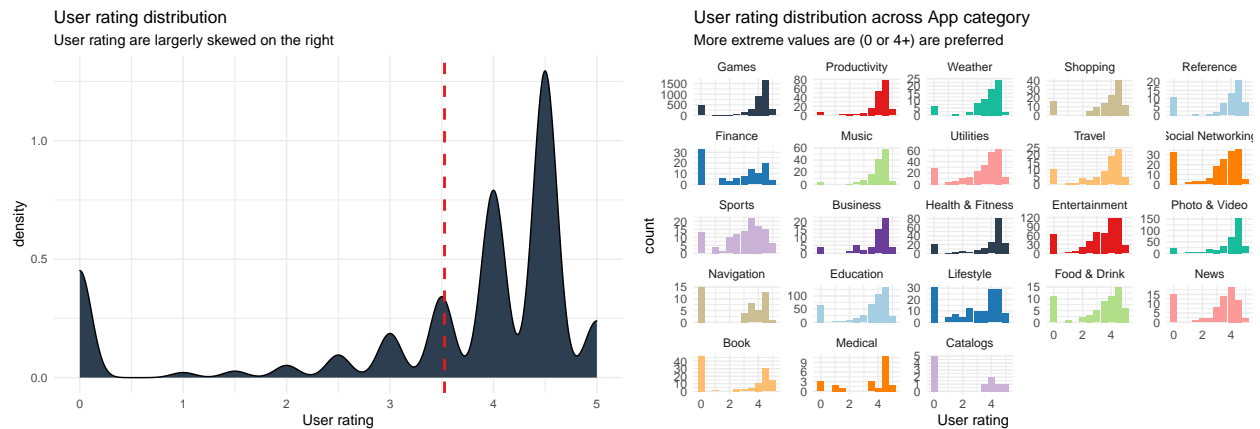
An estimation of the profits generated by the Apps was performed. The estimation was calculated multiplying the price by the total number of user rates, this can give as a bottom estimation of their revenues. According to the above, business Apps are the most profitable on average because even if their presence in the App Store is very small, they are really requested and their average price, as showed before, is really high. If we look at the top 10 performer in revenue is noticeable that the are almost all games. This shows why mobile gaming is a really hot topic right now.

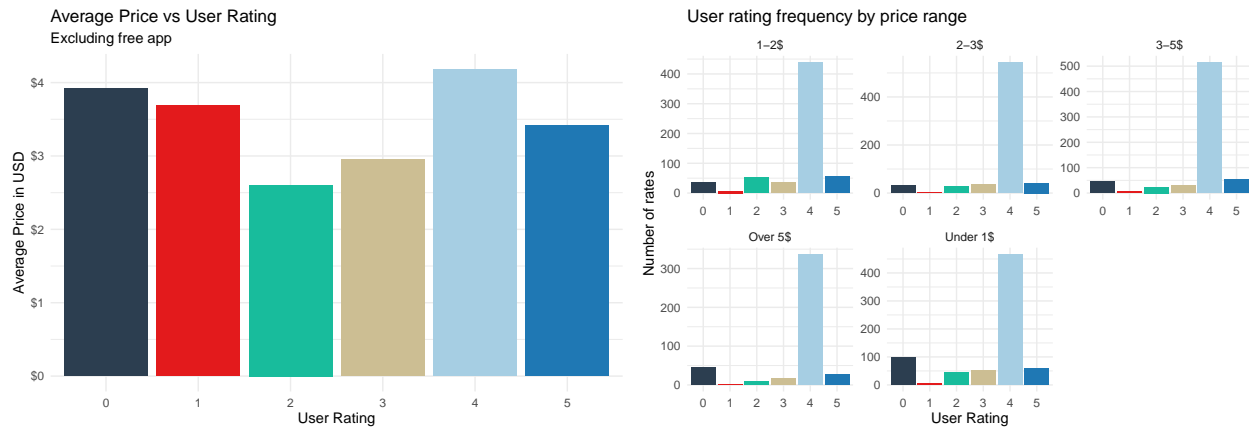
Estimated market bottom value by App category

Gaming App represents the biggest market share



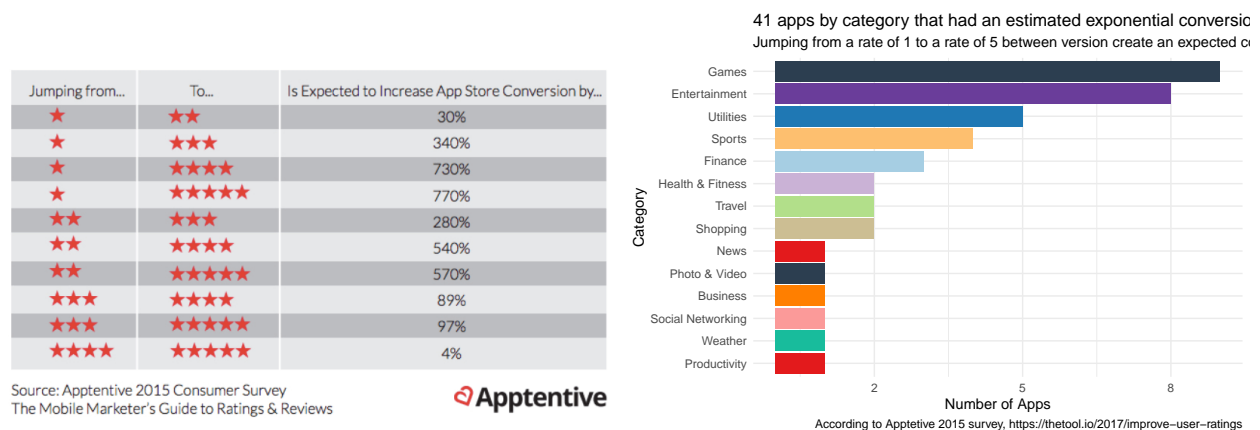
Even if Business Apps are the best in term of average revenues if all the revenues generated in the store are taken and split by category, Games is the one with the biggest share of generated revenues.





As well explained in this article by **Harvard Business Review** 5 stars review system has a big flaw. Users who have extreme experiences (either very bad or very good) are much more likely to leave feedback than users who have average experiences, thus creating selection biases, as showed in the first plot rates are concentrated on the extreme positive (4 or bigger) and negative (0). It is possible to notice the same effect across all the App category.

Price doesn't seem to affect the rating of the app generally even if we segment the price in 5 different ranges.



Thanks to Apptentive 2015 survey it was possible to estimate the effect of the change, if it happened, in rating between the previous and current version.